

Homework Assignment X - Extra Time Series Questions

Max H. Farrell – Chicago Booth
BUS41100 Applied Regression Analysis

NEVER DUE. Just for practice.

1 United States Gas Prices

In this question we will attempt to capture the time series dependence of weekly US gas prices. The data file `USGasPrice.csv` contains the `year`, `week`, and the `price` of gas from the eighth week of 1990 through the twenty-sixth week of 2003.

In each part below, we will use one or more of the tools we learned in class. For each part, plot the **(1)** fitted values along with the original values as well as **(2)** the ACF (see the slides for week 8).

- (a) As a first attempt, use annual frequencies to try to capture the idea that driving (and hence gasoline demand) is highly seasonal. Use the two plots asked for above to describe what goes wrong with this approach intuitively.
- (b) There appear to be several structural changes in the data (e.g. more recent years behave differently). Identify a reasonable set of time points when the series pattern changes dramatically and permanently and then fit a constant between each time point. What does this capture?
- (c) For each period you identified above, fit a separate annual cycle. What do you learn from the two plots ask for in the question?
- (d) Experiment with adding further cycles, time trends, and/or lagged values, in each case allowing each period you identified in **(d)** to have a separate fit. Try a “kitchen-sink” that includes cycles, trends, and lags.
- (e) Finally, what of the “kitchen-sink” model can you do without? That is, give the simplest model you can that still does nearly as well at capturing the time series dependence. What do you learn from this?

2 United Kingdom Gas Consumption

The goals of this question are to develop the best possible model for prediction of quarterly UK gas consumption (file `UKGasConsumption.csv`). Since one would expect gas consumption to increase with both population (a measure of personal consumption) and GDP (a measure of commercial production), the data consist of quarterly UK gas consumption (in millions of therms), inflation adjusted GDP, and population estimates for the years 1960 to 1986 (1987 for `GDP` and `pop`).

- (a) Find a regression model that best explains the time series for gas consumption (transform to log scale). You may incorporate the effect of GDP and population into your time series (again, consider a possible transformation of these variables).
- (b) Comment on your chosen model. For example, is there evidence of either mean reversion or a linear time trend in your series? What is the effect of the covariates on gas consumption?

Modeling Note: This data exhibits a common trait for quarterly data: autoregression on an annual basis. That means you want to include Y_{t-4} as your AR term, rather than the usual Y_{t-1} . For example,

```
gasdata <- read.csv("UKGasConsumption.csv")
QR <- 5:108
cos4 <- cos(QR*pi/2)
sin4 <- sin(QR*pi/2)
loggas <- log(gasdata$gas[QR])
loggaslast <- log(gasdata$gas[QR-4])
```

This provides periodic, AR, and linear effect variables.